



Title	Contrasting the effects of duration and number of syllables on the perceptual normalization of lexical tones
Author(s)	Ciocca, V; Francis, AL; Yau, TSK
Citation	147th Meeting of the Acoustical Society of America, New York, New York, USA, 24-28 May 2004, v. 115 n. 5 Pt 2, p. 2544
Issued Date	2004
URL	http://hdl.handle.net/10722/54391
Rights	Creative Commons: Attribution 3.0 Hong Kong License

the flows of the vocal tract. Acoustic information extracted from the speech waveform is mapped into inputs for HLSyn. This analysis by synthesis approach is a method to develop a more precise picture of the planning stage during speech production where the acoustic phonetics must be carefully planned and modified to achieve the correct target sounds. [Work funded by a grant provided by NIH.]

4aSC27. Text-to-phonemic transcription and parsing into mono-syllables of English text. Yugal Jusgir Mullick, S. S. Agrawal, Smita Tayal, and Manisha Goswami (CSIO, CSIRComplex, 2nd Fl., Pusa Campus, New Delhi, India)

The present paper describes a program that converts the English text (entered through the normal computer keyboard) into its phonemic representation and then parses it into mono-syllables. For every letter a set of context based rules is defined in lexical order. A default rule is also defined separately for each letter. Beginning from the first letter of the word the rules are checked and the most appropriate rule is applied on the letter to find its actual orthographic representation. If no matching rule is found, then the default rule is applied. Current rule sets the next position to be analyzed. Proceeding in the same manner orthographic representation for each word can be found. For example, “reading” is represented as “rEdiNX” by applying the following rules:

r→r move 1 position ahead
ead→Ed move 3 position ahead
i→i move 1 position ahead
ng→NX move 2 position ahead, i.e., end of word.

The phonemic representations obtained from the above procedure are parsed to get mono-syllabic representation for various combinations such as CVC, CVCC, CV, CVCVC, etc. For example, the above phonemic representation will be parsed as rEdiNX→/rE/ /diNX/. This study is a part of developing TTS for Indian English.

4aSC28. Signature of prosody in tonal realization: Evidence from Standard Chinese. Yiya Chen (Theoretical and Appl. Linguist. Dept., Edinburgh Univ., Adam Ferguson Bldg., EH8 9LL, UK)

It is by now widely accepted that the articulation of speech is influenced by the prosodic structure into which the utterance is organized. Furthermore, the effect of prosody on F_0 realization has been shown to be mainly phonological [Beckman and Pierrehumbert (1986); Selkirk and Shen (1990)]. This paper presents data from the F_0 realizations of lexical tones in Standard Chinese and shows that prosodic factors may influence the articulation of a lexical tone and induce phonetic variations in its surface F_0 contours, similar to the phonetic effect of prosody on segment articulation [de Jong (1995); Keating and Foureron (1997)]. Data were elicited from four native speakers of Standard Chinese producing all four lexical tones in different tonal contexts and under various focus conditions (i.e., under focus, no focus, and post focus), with three renditions for each condition. The observed F_0 variations are argued to be best analyzed as resulted from prosodically driven differences in the phonetic implementation of the lexical tonal targets, which in turn is induced by pragmatically driven differences in how distinctive an underlying tonal target should be realized. Implications of this study on the phonetic implementation of phonological tonal targets will also be discussed.

4aSC29. Variation of Taiwanese tones in conversation. Pei-Yu Hsieh and Jane Tsay (Inst. of Linguist., Chung Cheng Univ., 160 San-Xing, Min-Xiong, Chia-Yi 621, Taiwan, ROC, g9015001@ccu.edu.tw)

In laboratory research on tonal coarticulation in Taiwanese, one study [H.-B. Lin, Ph.D. dissertation, Univ. of Connecticut (1988)] reported a perseveratory effect but no anticipatory effect, while another [S.-H. Peng, *J. Phonetics* 25, 371–400 (1997)] found a significant anticipatory effect.

Peng also found tonal variation due to prosodic positions. Unlike these previous laboratory studies, this study attempts to investigate tonal coarticulation and prosodic effects on Taiwanese tones using natural conversations from the Taiwanese Spoken Corpus (Tsay and Myers, 2004), of which 56 min of recorded conversations were analyzed. Consistent with Lin, the results showed that tone is more affected by the preceding tone than by the following tone. The slope is more influenced by the preceding tone as well. As for prosodic effects, the results confirmed Peng, showing that F_0 is the lowest in utterance-final position, while in other phrase-final positions it is slightly lower than in non-phrase-final position. This study thus demonstrates the results obtained in the laboratory do indeed carry over into actual conversation.

4aSC30. Contrasting the effects of duration and number of syllables on the perceptual normalization of lexical tones. Valter Ciocca, Alexander L. Francis, and Teresa S.-K. Yau (Div. of Speech & Hearing Sci., Univ. of Hong Kong, 5/F, 34 Hospital Rd., Hong Kong, HKSAR, China, vciocca@hkusua.hku.hk)

In tonal languages, syllabic fundamental frequency (F_0) patterns (“lexical tones”) convey lexical meaning. Listeners need to relate such pitch patterns to the pitch range of a speaker (“tone normalization”) to accurately identify lexical tones. This study investigated the amount of tonal information required to perform tone normalization. A target CV syllable, perceived as either a high level, a low level, or a mid level Cantonese tone, was preceded by a four-syllable carrier sentence whose F_0 was shifted (1 semitone), or not shifted. Four conditions were obtained by gating one, two, three, or four syllables from the onset of the target. Presentation rate (normal versus fast) was set such that the duration of the one, two, and three syllable conditions (normal carrier) was equal to that of the two, three, and four syllable conditions (fast carrier). Results suggest that tone normalization is largely accomplished within 250 ms or so prior to target onset, independent of the number of syllables; additional tonal information produces a relatively small increase in tone normalization. Implications for models of lexical tone normalization will be discussed. [Work supported by the RGC of the Hong Kong SAR, Project No. HKU 7193/00H.]

4aSC31. Effects of native language experience on perceptual learning of Cantonese lexical tones. Alexander L. Francis (Dept. of Audiol. and Speech Sci., Purdue Univ., Heavilon Hall, 500 Oval Dr., West Lafayette, IN 47907, francisa@purdue.edu), Valter Ciocca (The Univ. of Hong Kong, Hong Kong, China), and Lian Ma (Beijing Medical Univ., Beijing, China)

In a tonal language syllabic pitch patterns contribute to lexical meaning. Perceptual assimilation models of cross-language perception predict speakers of another tonal language should assimilate Cantonese lexical tones to native tonal categories, affecting identification, discrimination and acquisition. For nontonal language speakers, two possibilities exist. If pitch information is ignored, vowels with different tones should assimilate to the same native category, lowering performance. If tonal information is attended but unused in native categorization, Cantonese tones could be nonassimilable and therefore easily discriminated, and possibly easily identified or learned. Here, native speakers of Mandarin Chinese and American English were trained to identify Cantonese words differing in lexical tone. Discrimination and identification were tested before and after training. Both groups initially performed well on upper register tones (high level, high rising, mid level) and poorly on lower (low falling, low level, low rising). Mandarin listeners improved most at identifying low falling tones; English listeners improved most on low level and low rising tones. Training primarily appeared to improve listeners’ ability to make categorical decisions based on direction of pitch change, a feature reportedly under-attended by English speakers, but preferred by Mandarin speakers. [Work supported by research funding from The University of Hong Kong.]